RNA (2001), *7*:421–434. Cambridge University Press. Printed in the USA.

The sequence complementarity between HIV-1 5′ splice site SD4 and U1 snRNA determines the steady-state level of an unstable *env* pre-mRNA

SUSANNE KAMMLER,[1] CORDULA LEURS,[1] MARCEL FREUND,[1] JÖRG KRUMMHEUER,[1] KERSTIN SEIDEL,[1] THOMAS Ø. TANGE,[2] METTE K. LUND,[2] JØRGEN KJEMS,[2] ANDREAS SCHEID,[1] and HEINER SCHAAL[1]

[1]Institut für Medizinische Mikrobiologie und Virologie, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany
[2]Department of Molecular and Structural Biology, University of Aarhus, DK-8000 Aarhus C, Denmark

ABSTRACT

HIV-1 *env* expression from certain subgenomic vectors requires the viral regulatory protein Rev, its target sequence RRE, and a 5′ splice site upstream of the *env* open reading frame. To determine the role of this splice site in the 5′-splice-site-dependent Rev-mediated *env* gene expression, we have subjected the HIV-1 5′ splice site, SD4, to a mutational analysis and have analyzed the effect of those mutations on *env* expression. The results demonstrate that the overall strength of hydrogen bonding between the 5′ splice site, SD4, and the free 5′ end of the U1 snRNA correlates with *env* expression efficiency, as long as *env* expression is suboptimal, and that a continuous stretch of 14 hydrogen bonds can lead to full *env* expression, as a result of stabilizing the pre-mRNA. The U1 snRNA-mediated stabilization is independent of functional splicing, as a mismatch in position +1 of the 5′ splice site that led to loss of detectable amounts of spliced transcripts did not preclude stabilization and expression of the unspliced *env* mRNA, provided that Rev enables its nuclear export. The nucleotides capable of participating in U1 snRNA:pre-mRNA interaction include positions –3 to +8 of the 5′ splice site and all 11 nt constituting the single-stranded 5′ end of U1 snRNA. Moreover, *env* gene expression is significantly decreased upon the introduction of point mutations in several upstream GAR nucleotide motifs, which are mediating SF2/ASF responsiveness in an in vitro splicing assay. This suggests that the GAR sequences may play a role in stabilizing the pre-mRNA by sequestering U1 snRNP to SD4.

Keywords: envelope; ESE; HIV-1; RNA splicing; RNA stability; U1 snRNA

INTRODUCTION

Most eukaryotic primary transcripts contain introns that have to be removed prior to their nucleocytoplasmic export to ensure translation of a continuous open reading frame (ORF). Splicing takes place in two catalytic steps within a large multicomponent complex, the spliceosome (for reviews, see Moore et al., 1993; Burge et al., 1998; Reed, 2000). One of the earliest steps in spliceosome assembly is the duplex formation between the free 5′ end of the U1 snRNA and the 5′ splice site. The RNA duplex is not only stabilized through its hydrogen bonding but almost certainly through additional interactions of protein components with the pre-mRNA also in the vicinity of the 5′ splice site (Puig et al., 1999; Zhang & Rosbash, 1999; Del Gatto-Konczak et al., 2000). Beside its role in the splicing reaction, the integrity of a 5′ splice site is important for nuclear pre-mRNA stability as has been shown by point mutations within the 5′ splice site of, for example, the HIV-1 *tat*/*rev* intron (Lu et al., 1990) or polyoma virus late transcripts (Barrett et al., 1995) leading to a decrease in the accumulation of unspliced transcripts. Moreover, binding of snRNPs leads to nuclear retention of the RNA and it is generally recognized that completion of the splicing reaction removes this obstacle to RNA nucleocytoplasmic export (Chang & Sharp, 1989; Legrain & Rosbash, 1989; Hamm & Mattaj, 1990; Huang & Carmichael, 1996).

There remains a conceptual problem of how the nucleocytoplasmic export of unspliced RNA equipped with suitable snRNP binding sites occurs. This is the case for the nucleocytoplasmic export of a variety of retro- and pararetrovirus mRNAs (Cullen, 1992; Kiss-Laszlo

Reprint requests to: Heiner Schaal, Institut für Medizinische Mikrobiologie und Virologie, Geb. 22.21, Universitätsstr. 1, D-40225 Düsseldorf, Germany; e-mail: schaal@uni-duesseldorf.de.

& Hohn, 1996; Bodem et al., 1997). The overlapping of their genes entails that one and the same sequence may be part of an intron or of an exon, depending on which gene is being expressed. To enable the use of their full genomic potential, the excision of introns has to be restrained by regulatory processes, for instance, through inefficient 3′ splice sites (Katz & Skalka, 1990; Fu et al., 1991; McNally & Beemon, 1992; Staffa & Cochrane, 1994; Dyhr-Mikkelsen & Kjems, 1995; O'Reilly et al., 1995; Zhang & Stoltzfus, 1995; Si et al., 1997) or *cis*-acting splicing enhancer/silencer elements (Amendt et al., 1995; Staffa & Cochrane, 1995; Wentz et al., 1997; Si et al., 1998; Caputi et al., 1999). In the case of HIV-1, the predominant cytoplasmic appearance of the unspliced *env* mRNA, which contains the *tat*/*rev* intron (Fig. 1), is regulated by the viral protein Rev through interaction with its RNA target sequence the Rev responsive element (RRE; for a review see Emerman & Malim, 1998; Hope, 1999). By utilizing a non-mRNA export pathway that is coupled to the Ran-GTPase cycle, Rev counteracts the splicing pathway and nuclear retention (Fischer et al., 1999).

Rev-mediated *env* expression was originally suggested to inhibit splicing directly depending on inefficient splice signals (Chang & Sharp, 1989), which would also imply a less efficient nuclear retention signal. However, in HIV-1, the 5′ splice site SD4 of the second intron of the *tat*/*rev* transcript, that is, upstream of the *env* ORF, turned out to be highly efficient, whereas the 3′ splice site, SA7, is suboptimal (McNally & Beemon, 1992; Staffa & Cochrane, 1994; Dyhr-Mikkelsen & Kjems, 1995; O'Reilly et al., 1995; Si et al., 1998). Lu et al. (1990) have conclusively shown that the cytoplasmic expression of an SV40 late promoter driven 3′ subgenomic HIV-1 fragment requires that the 5′ splice site be effectively recognized by U1 snRNA and that binding of U1 snRNA affects the steady-state level of *env* mRNA. In other experimental systems, deletion of the 5′ splice site did not abolish Rev-mediated cytoplasmic appearance of RRE-containing RNAs (Emerman et al., 1989; Felber et al., 1989; Fischer et al., 1994, 1999; Nasioulas et al., 1994). This suggested to us that the splice site also serves an RNA protective function for the 5′-splice-site-dependent Rev-mediated *env* gene expression.

A starting point for the design of our experiments was the observation by Lu et al. (1990) that a mutation in position +5 of SD4 prevented *env* expression and could be compensated by a mutation in the corresponding nucleotide of the U1 snRNA. The present analysis does not focus primarily on compensatory effects in U1 snRNA, but on the nucleotides within SD4 and the extent of base-pairing to U1 snRNA. Here, we demonstrate that the sequence of the HIV-1 5′ splice site SD4 and of an upstream exon splicing enhancer (ESE) has profound and predictable effects on the steady-state poly(A)$^+$ RNA levels expressed from vectors in which SD4 is the 5′ splice site closest to the cap. Moreover, we show that the U1 snRNP association stabilizes the pre-mRNA in a process independent of splicing.

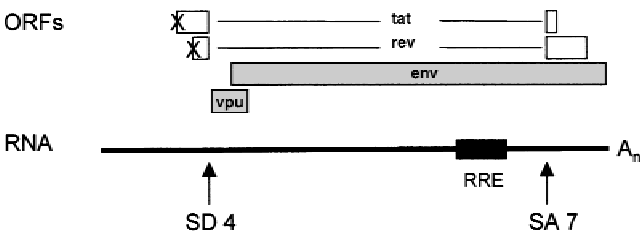

FIGURE 1. Schematic drawing of the positions of the 5′ and 3′ splice sites flanking the *tat*/*rev* intron of the HIV-1 transcript synthesized from the *env* expression vector SV E/X tat$^-$ rev$^-$. The *Eco*RI-*Xho*I fragment of pNLA1 (Strebel et al., 1987) was cloned into the SV40 early expression vector pSVT7 (Bird et al., 1987). ORFs indicated as open boxes are translationally silent due to their mutated translational start codons "X". Based on SV E/X tat$^-$ rev$^-$, the mutant vectors are named according the position of the 5′ splice site that is mutated and the substituted nucleotide. The first nucleotide of the intron is defined as +1. In SD4$^-$, only the 3 nt of the *tat* exon 1 stop codon are complementary to the U1 snRNA.

RESULTS

Hydrogen bonding between the HIV-1 5′ splice site SD4 and U1 snRNA determines Env expression

To investigate the role of a 5′ splice site for Rev function, we analyzed the Rev-dependent *env* expression of an SV40-based *env* expression vector containing mutations in SD4 (Fig. 1). The *env* expression vector SV E/X tat$^-$ rev$^-$ used in our study (Fig. 1) contains the *Eco*RI-*Xho*I fragment of pNLA1 (Strebel et al., 1987) encompassing the downstream half of the HIV-1 genomic sequence. The vector possesses the expression potential for Vpu and Env, but not for Tat and Rev because of point mutations in their translational start codons. The mutation of the *tat* AUG was introduced to exclude the translational inhibition of the downstream *vpu* and *env* AUGs by the strong *tat* AUG (Schwartz et al., 1990). The mutation of the *rev* AUG was designed to exclude variable *rev* expression as a cause for different efficiency of *env* expression. Equal Rev levels were provided by cotransfection with a Rev expression vector. All transfection experiments were performed in HeLa cells under nonreplicative conditions.

To determine the minimal number of neighboring hydrogen bonds between SD4 and U1 snRNA that are necessary for *env* expression, we analyzed the parent construct with the authentic SD4 sequence and two mutant SD4 sequences, containing mismatches in position +7 (7C) and +6/+7 (6A7C) (Fig. 2). This reduces the complementary bases to U1 snRNA from 7 in SD4 to 6 and 5 in the mutant plasmids and the hydrogen bonds from 16 in SD4 to 14 and 12, respec-

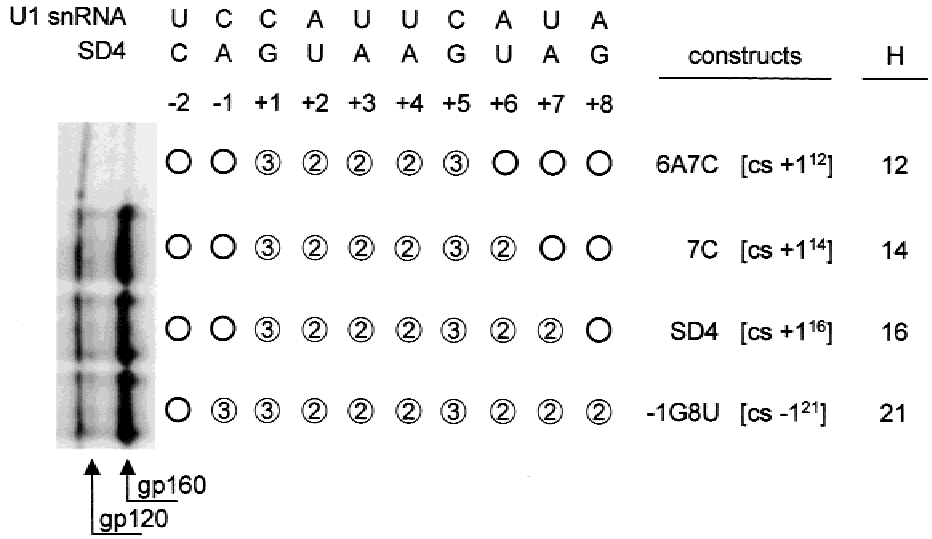

FIGURE 2. Immunoblot analysis (left) of glycoproteins expressed in HeLa cells transfected with *env* expression vectors carrying mutations of the *tat*/*rev* 5′ splice site, SVcrev, and pGL3-control. For each plasmid, transfection using the transfection reagent FuGene^{TM}6 was done with each 1 μg of the *env* expression vector, 1 μg SVcrev, and 0.5 μg pGL3-control. The mutations of the *tat*/*rev* 5′ splice site are depicted on the right, given the position and the substituted nucleotide. A synonymous nomenclature with respect to the continuous stretch (cs) of formed hydrogen bonds starting at the position of the cs on SD4 and the number of hydrogen bonds indicated by the superscript is given in brackets. Column H states the number of predicted hydrogen bonds. In the sketch in the middle, the hydrogen bonds that are predicted to form between the *tat*/*rev* 5′ splice site and U1 snRNA are shown as circles (mismatch: open circle; 2: two hydrogen bonds; 3: three hydrogen bonds).

tively. In addition, we analyzed a construct containing mutations in positions -1 and $+8$ ($-$1G8U) in SD4, increasing the complementary bases to nine and the hydrogen bonds to 21. Immunoblot analysis of cells transfected with these constructs showed wild-type *env* expression from the 7C and $-$1G8U constructs, but severely affected expression from mutant 6A7C. This implies that the minimal number of adjacent base pairings is six, and that the number of adjacent hydrogen bonds necessary for wild-type *env* expression is greater than 12.

To determine the position of the hydrogen bonds enabling *env* expression, we constructed additional 5′ splice site mutants with a continuous stretch (cs mutants) of 14 to 16 potential hydrogen bonds with U1 snRNA, shifted towards the 5′ direction of the splice site (Fig. 3A, cf. also Fig. 3B). Transfection experiments with these mutants show that the positions -2 (cf. mutant cs -2^{14}, i.e., base pairing extended to splice site position -2, 14 predicted hydrogen bonds) and -3 (cf. mutant cs -3^{15}) are capable of compensating for mismatches in positions $+5$ and $+6$, indicating that they can participate in U1 base pairing. In contrast, mutant cs -4^{16} has lost its ability to express *env*, possibly due to a lack of base pairing in position -4, resulting in a reduction of hydrogen bonds from 16 to 13. This was confirmed by efficient *env* expression from the mutant with one match added to the stretch -4^{16} in position $+3$ (cs -4^{18}), but loss of *env* expression by introducing a mismatch at position -4 (cs -3^{13}). Taken together, these findings confirm that six is the minimum number of adjacent base pairings and 14 is the minimum number of hydrogen bonds that is required for glycoprotein expression in our experimental system.

Central mismatches in the SD4 sequence complementary to U1 snRNA affect *env* expression

To determine the requirement for *env* expression of a discontinuous stretch of hydrogen bonds, we introduced point mutations in the center of the SD4 sequence at positions $+3$, $+4$, or $+5$. In these mutants, *env* expression is obliterated (Fig. 4A), despite the fact that the total number of hydrogen bonds in the mutants 3U and 4C equals the minimum of 14 as determined for a continuous stretch of hydrogen bonds (Figs. 2 and 3). Thus, in the case of discontinuous stretches of hybridization potential, more than 14 hydrogen bonds are required. Introduction of additional mutations to allow hydrogen bonding with U1 snRNA at positions $+8$ (Fig. 4B) or -1 (Fig. 4C) or both (Fig. 4D) can in some instances compensate for the mismatch. The suboptimal hydrogen bonding of mutant 4C (Fig. 4A) with seven hydrogen bonds on either side of the mismatch, can be compensated by prolonging the base-pairing region by 1 nt on either the 5′ or 3′ side (mutant 4C8U, Fig. 4B, and $-$1G4C, Fig. 4C). With fewer than seven hydrogen bonds on one side of the mismatch, for example, in mutant 5U, additional hydrogen bonds

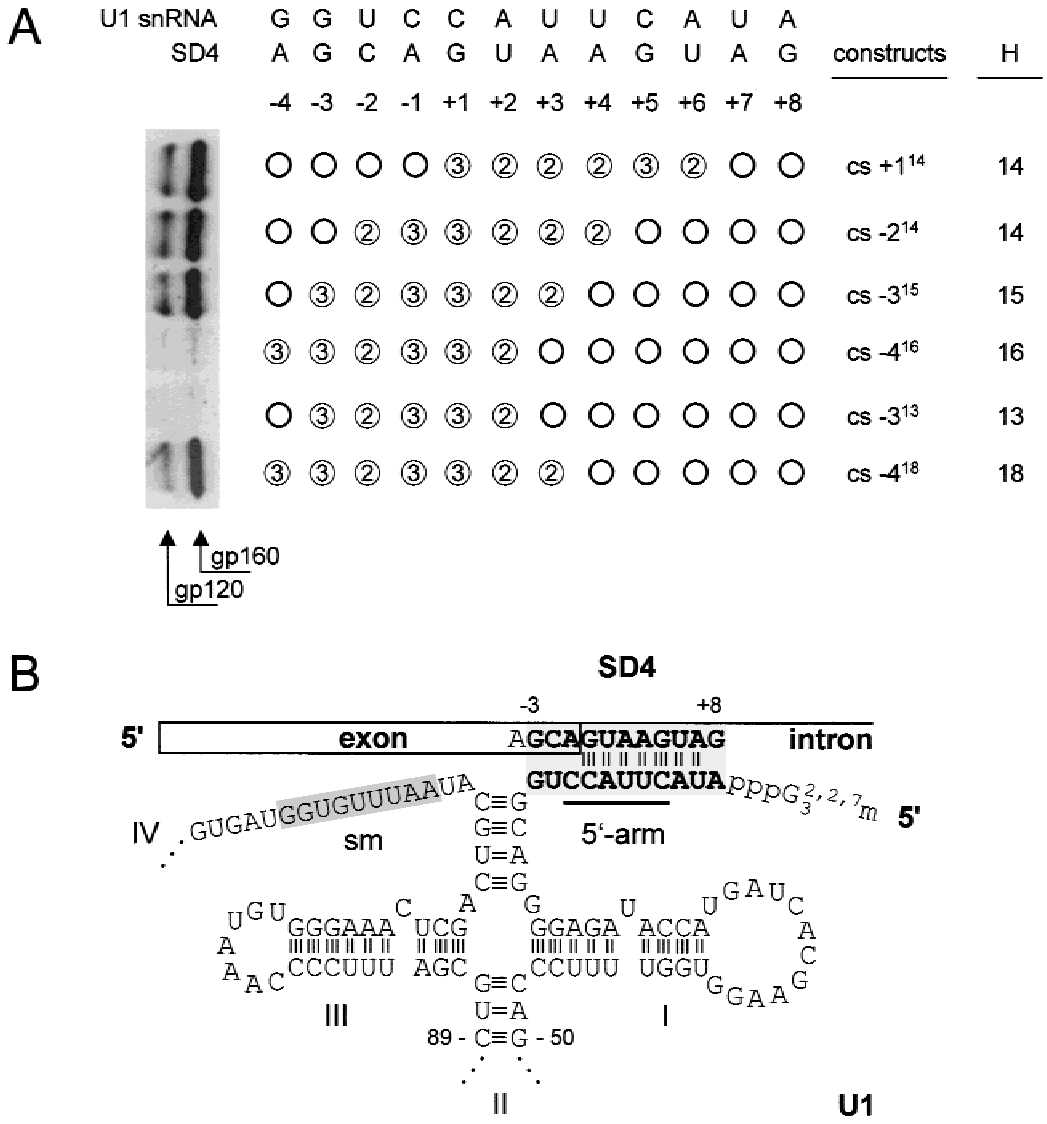

FIGURE 3. **A**: Immunoblot analysis (left) of glycoproteins expressed in HeLa cells transfected with SVcrev, pGL3-control, and env expression vectors carrying mutations in the *tat*/*rev* 5′ splice site. See Figure 2 for an explanation. **B**: Hybridization of sequences of the *tat*/*rev* exon/intron border with the 5′ arm of U1 snRNA (according to Branlant et al., 1981).

on both the 5′ and the 3′ side are necessary to restore *env* expression (−1G5U8U, Fig. 4D). In conclusion, 6–12 hydrogen bonds on either side of the mismatch are necessary for *env* expression, and the exact number of hydrogen bonds on each side depends on the strength of the base pairing on the opposite side of the mismatch.

To prove that the observed correlation between *env* expression and the potential number of hydrogen bonds is due to the RNA duplex formation between the U1 snRNA and SD4, we repeated the transfection experiment with some of the 3U mutants (cf. Fig. 5, lanes wt U1 and 3U mutants in Fig. 4A–D), but with coexpression of a suppressor U1 snRNA matching the 3U mutation (Fig. 5, lanes 6A U1). As predicted, *env* expression was increased for all 3U mutants when this suppressor U1 snRNA was coexpressed. Moreover, for the two mutants with five hydrogen bonds on one side of the mismatch (mutants 3U and 3U8U), but different numbers of hydrogen bonds on the opposite side, *env* expression increased with an increased total number of hydrogen bonds. On the contrary, for the other two mutants that have eight instead of five hydrogen bonds on one side of the 3U mismatch (mutants −1G3U and −1G3U8U), the correlation between *env* expression and total number of hydrogen bonds disappears. This shows that maximal *env* expression can already be obtained when forming an RNA duplex with a less than maximum number of hydrogen bonds. Thus, the contribution of any single nucleotide, for example, position +8, to the formation of the RNA duplex only becomes important if *env* expression is suboptimal.

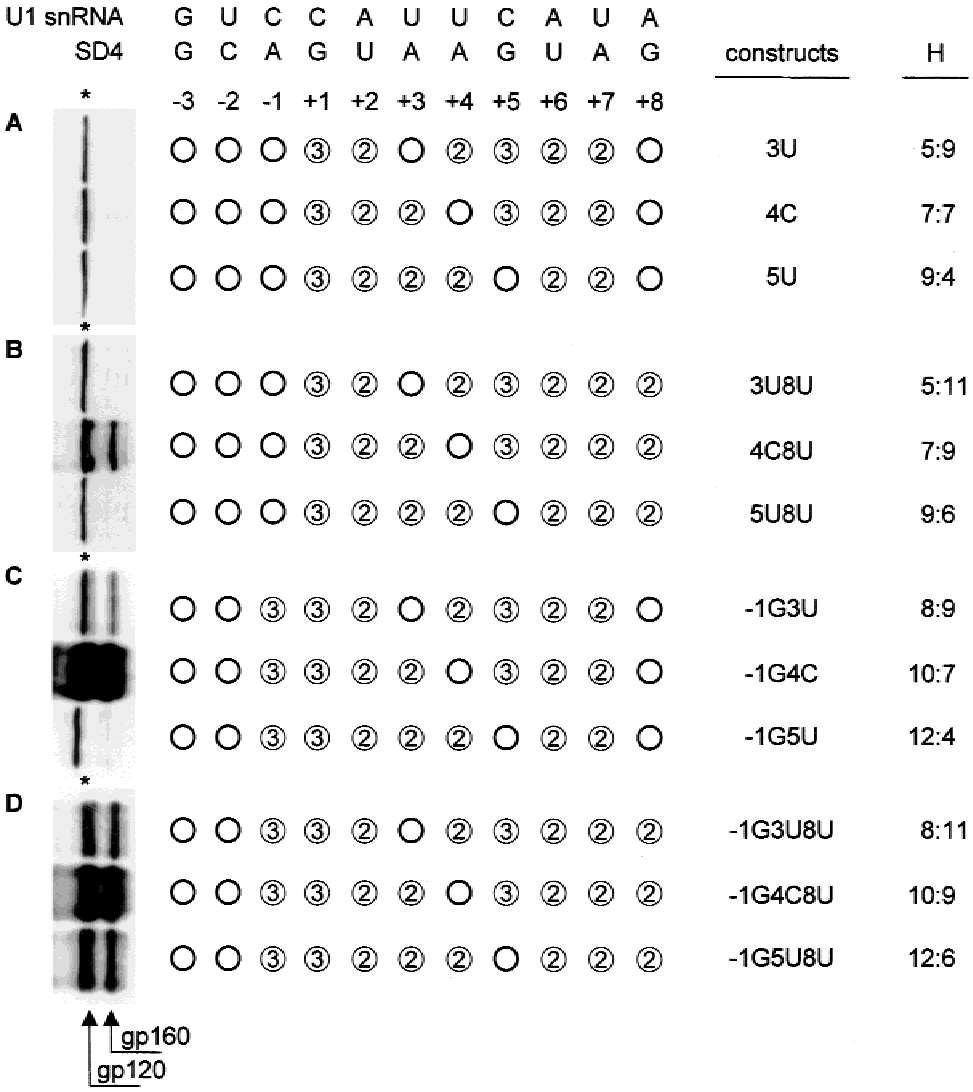

FIGURE 4. Immunoblot analysis (left) of glycoproteins expressed in HeLa cells transfected with SVcrev, pGL3-control, and *env* expression vectors carrying mutations in the *tat*/*rev* 5′ splice site. See Figure 2 for an explanation. A cellular protein cross reacting with the mAb is marked with an asterisk. Column H: Number of hydrogen bonds at either side of the unpaired nucleotide.

Beyond identification of the hydrogen bonding patterns suitable for *env* expression, our data indicate that a mismatch in position +4 is more readily compensated by extending the flanking base pairing than a mismatch in position +3 or +5 (Fig. 4). To analyze whether this is due to a favorable distribution of hydrogen bonds in the mutant 4C, or to a preference for hydrogen bonding of position +3 as compared to position +4, we constructed two additional 3U mutants with an arrangement of hydrogen bonds that is similar to that of the 4C mutants (Fig. 6). The immunoblot analysis of these constructs shows that the 4C mutant with seven hydrogen bonds at the 5′ side and nine at the 3′ side of the mismatch expresses higher amounts of Env than the 3U mutant with eight and nine hydrogen bonds (Fig. 6, cf. mutants 4C8U and −1G3U). The same result could be obtained in the comparison of mutants with 10 and 7 hydrogen bonds on the 5′ and 3′ sides of the mismatch (Fig. 6, cf. −1G4C and −2A−1G3U7U). In all cases, the mismatch in position +4 is more readily compensated than a mismatch in position +3. This finding suggests that the position of the mismatch or the positions of the compensatory matches also exert an influence on *env* expression, albeit to a lesser extent than the total number of hydrogen bonds. Additionally, comparison of two mismatch mutants, both of which can form an RNA duplex with a discontinuous, but identical total number of 17 hydrogen bonds (mutants −1G3U and −2A−1G3U7U), suggests that the length of a single continuous stretch of hydrogen bonds is the primary determinant for *env* expression the closer it approximates the critical num-

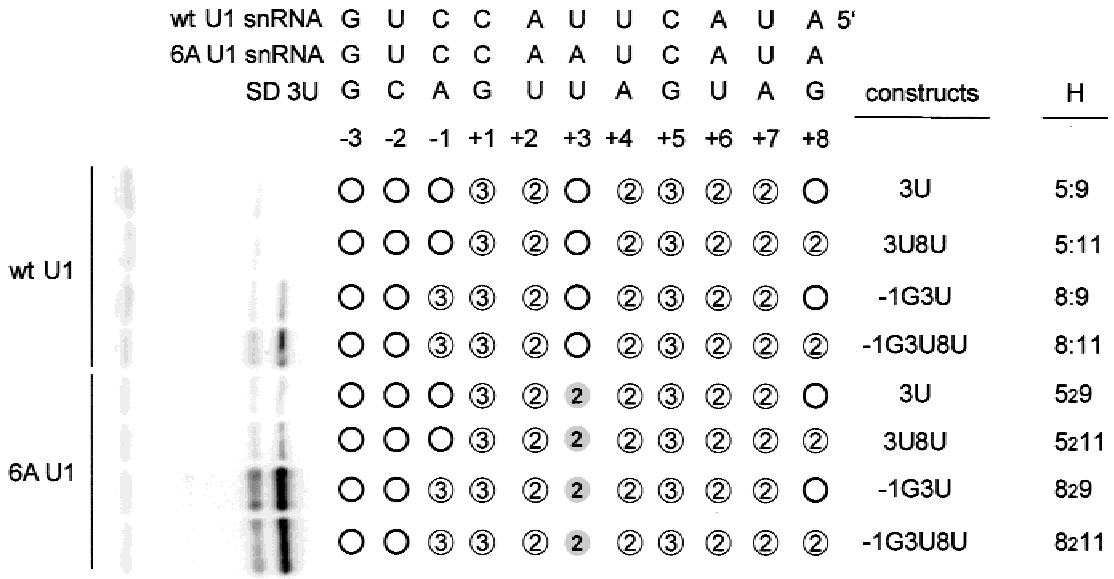

FIGURE 5. Immunoblot analysis (left) of glycoproteins expressed in HeLa-T4$^+$ cells transfected with SVcrev, pGL3-control, *env* expression vectors carrying mutations in the *tat*/*rev* 5′ splice site, and the expression plasmid coding for wt U1 snRNA (wt U1) or 6A U1 snRNA (6A U1). See Figure 2 for an explanation. Column H: Number of hydrogen bonds at either side of the unpaired nucleotide.

ber of hydrogen bonds that, in itself, would be sufficient for *env* expression (Fig. 6, cf. mutants −1G3U and −2A−1G3U7U).

Complementarity of the 5′ splice site to U1 snRNA is essential for RNA stability

To evaluate the splicing efficiency of the mutants in vivo, we performed northern blot analyses of transient transfected HeLa cells in the presence and absence of Rev (Figs. 7 and 8).

Analysis of the mutants containing continuous stretch complementary to U1 snRNA shifted towards the 5′ end of the splice site (Fig. 3A) revealed that all of the mutants with more than 13 hydrogen bonds (cs + 1[14], cs − 2[14], cs − 3[15], cs − 4[18]; lanes 1–3 and 6) are spliced efficiently if the continuous stretch starts at position −3 or at a position further downstream (Fig. 7).

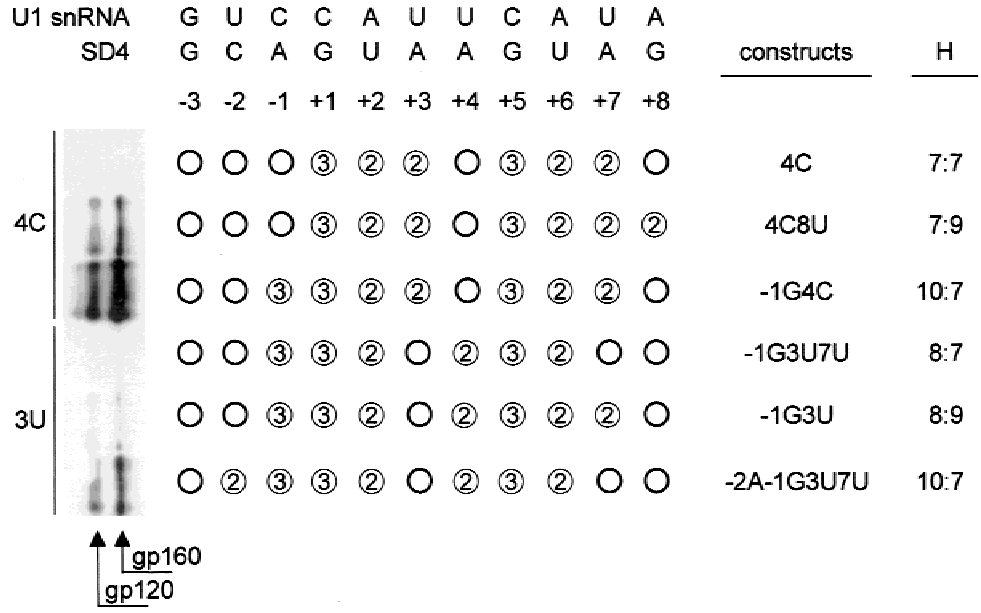

FIGURE 6. Immunoblot analysis (left) of glycoproteins expressed in HeLa cells transfected with SVcrev, pGL3-control, and *env* expression vectors carrying mutations in the *tat*/*rev* 5′ splice site. See Figure 2 for an explanation. Column H: Number of hydrogen bonds at either side of the unpaired nucleotide.

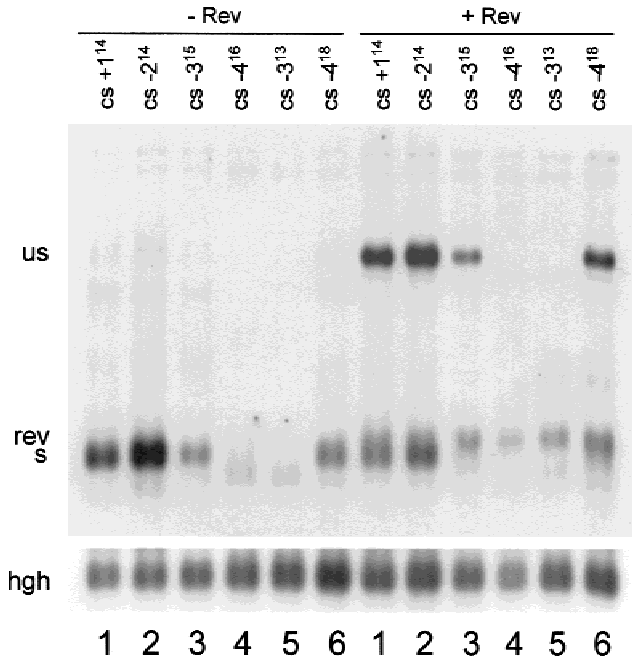

FIGURE 7. Northern blot analysis of poly(A)$^+$ RNA isolated from HeLa-T4$^+$ cells transfected with SV E/X tat$^-$ rev$^-$ constructs with a continuous stretch of complementarity to U1 snRNA starting at positions $+1$ to -4 (cf. Fig. 3) in the presence (+ Rev) and absence of Rev (− Rev). Lanes 1: SV E/X tat$^-$ rev$^-$ cs $+1^{14}$; lanes 2: SV E/X tat$^-$ rev$^-$ cs -2^{14}; lanes 3: SV E/X tat$^-$ rev$^-$ cs -3^{15}; lanes 4: SV E/X tat$^-$ rev$^-$ cs -4^{16}; lanes 5: SV E/X tat$^-$ rev$^-$ cs -3^{13}; lanes 6: SV E/X tat$^-$ rev$^-$ cs -4^{18}. us: unspliced transcript; s: spliced transcript; rev: cross-hybridization signal with the transcript of the co-transfected Rev expression plasmid; hgh: human growth hormone transcript visualizing transfection efficiency.

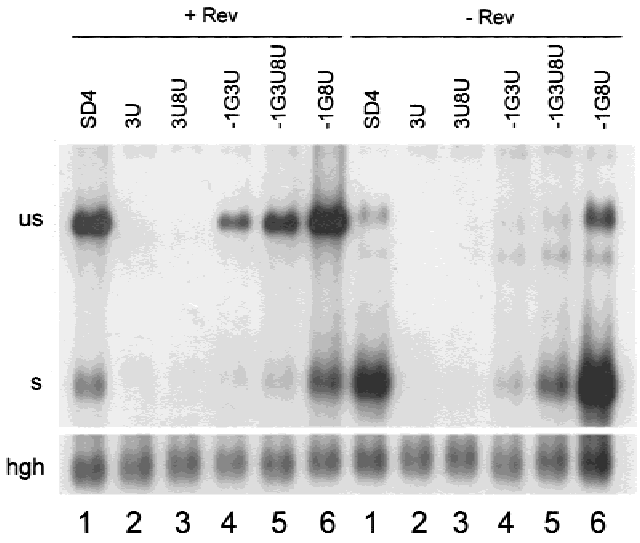

FIGURE 8. Northern blot analysis of poly(A)$^+$ RNA isolated from HeLa-T4$^+$ cells transfected with SV E/X tat$^-$ rev$^-$ constructs with a mismatch in the stretch of complementarity to U1 snRNA in position $+3$ (cf. Fig. 4) in the presence (+ Rev) and absence of Rev (− Rev). Lanes 1: SV E/X tat$^-$ rev$^-$; lanes 2: SV E/X tat$^-$ rev$^-$ 3U; lanes 3: SV E/X tat$^-$ rev$^-$ 3U8U; lanes 4: SV E/X tat$^-$ rev$^-$ −1G3U; lanes 5: SV E/X tat$^-$ rev$^-$ −1G3U8U; lanes 6: SV E/X tat$^-$ rev$^-$ −1G8U. us: unspliced transcript; s: spliced transcript; hgh: human growth hormone transcript.

However, the steady-state level of total poly(A)$^+$ mRNA decreases with those mutants whose continuous stretch of hydrogen bonding ends at position $+3$ (cs $-$ 3^{15} and cs $-$ 4^{18}; lanes 3 and 6), indicating that a continuous stretch of 15 hydrogen bonds, starting at position -3, would not be sufficient for an efficient stabilization. In the absence of Rev, the spliced RNA is detected whereas in the presence of Rev, most of the RNA is unspliced. Neither spliced nor unspliced RNA is detectable with cs $-$ 4^{16} and cs $-$ 3^{13} (lanes 4 and 5) obviating any conclusion regarding the minimal requirements of hydrogen bonds for the splice function in these unstable transcripts.

To analyze mutants with discontinuous complementarity to U1 snRNA, mutants derived from the 3U construct were chosen (Fig. 4). In the presence of Rev, the poly(A)$^+$ steady-state levels of unspliced *env* mRNAs correlate with Env formation (cf. Fig. 8, lanes 1–6 "+Rev", and Fig. 4), and the amount of spliced mRNA appears to be a constant proportion. Thus, the SD4 mutations strongly influence the poly(A)$^+$ steady-state levels of unspliced *env* mRNAs. In the absence of Rev, the spliced form of mRNA is dominating but the influence of the SD4 mutations on mRNA levels is comparable to those formed in the presence of Rev (Fig. 8, cf. lanes 1–6 "−Rev" and "+Rev"), suggesting that the stabilization of the RNA is independent of the splicing reaction. No subgenomic HIV-1 specific RNA was detected in cells transfected with mutants 3U (Fig. 8, lanes 2) and 3U8U (Fig. 8, lanes 3), neither in the presence nor absence of Rev (Fig. 8, lanes 2 and 3), implying that insufficient hydrogen bonding with U1 snRNA leads to nuclear degradation.

Similar results were also obtained with another two series of vectors that differed only in the presence of an RRE within the intron of their model transcription unit (data not shown). Furthermore, this indicates that the 5′-splice-site-mediated RNA stabilization mechanism is dependent neither on the RRE nor on the Rev/RRE interaction.

Nucleotide G in position +1 of the 5′ splice site is obligatory for the splice function, but not for RNA stabilization

To address the question of whether the splice function of the 5′ splice site is a prerequisite for the function as a stabilizing element, we dissected these two functions by including the position $+1$ in the mutational analysis of the SD4 sequence. Because the dinucleotide GT in position $+1$ and $+2$ is highly conserved in functional 5′ splice sites among the U2-type introns and because we have shown that at least 14 hydrogen bonds should be sufficient to stabilize the RNA, we mutated the guanine in position $+1$ to a cytosine but retained the potential for the formation of 13 to 23 hydrogen bonds to

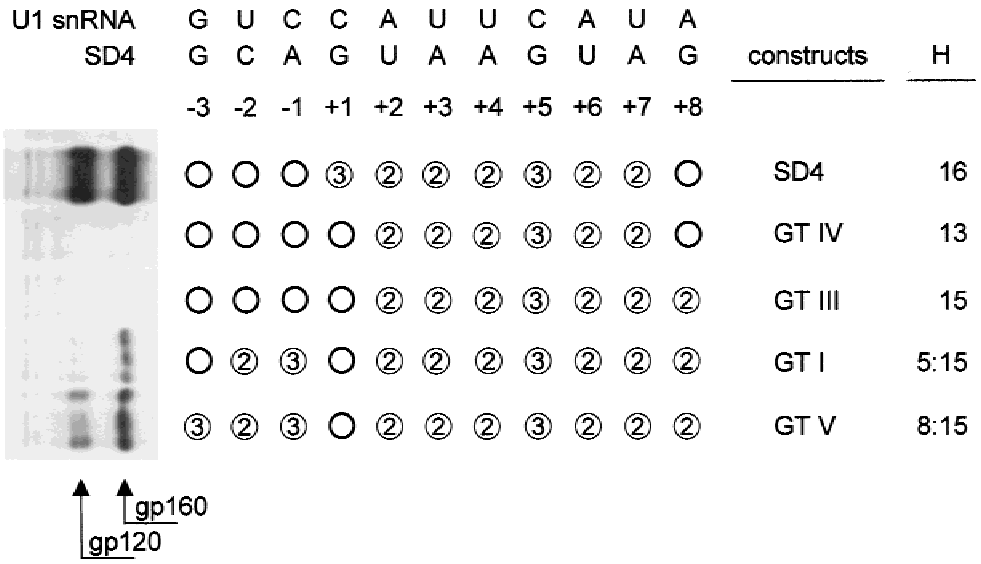

FIGURE 9. Immunoblot analysis (left) of glycoproteins expressed in HeLa cells transfected with SVcrev, pGL3-control, and *env* expression vectors carrying mutations in the *tat*/*rev* 5′ splice. See Figure 2 for an explanation. Column H: Number of hydrogen bonds at either side of the unpaired nucleotide.

U1 snRNA from position −3 to +8 (Fig. 9). These altered SD4 sequences were introduced into the *env* expression vector SV E/X tat$^-$ rev$^-$ and analyzed for their influence on Rev-dependent *env* expression and their splicing pattern. The immunoblot analysis (Fig. 9) shows that a mutation in position +1 in the context of extended flanking complementarity yielded moderate *env* expression (Fig. 9, mutants GTI and GTV), indicating that the G in position +1 is not obligatory for RNA stabilization.

The northern blot analysis of the corresponding RNAs (Fig. 10) shows that the levels of steady-state poly(A)$^+$ unspliced RNAs synthesized in the presence of Rev parallels the levels of Env expression shown in Figure 9. In the absence of Rev, spliced RNA is only detected for the wild-type SD4 (Fig. 10, lane 1), but not in any of the constructs containing a mutation at position +1 (Fig. 10, lanes 2–5 "−Rev").

These results imply that the presence of G in position +1 of the 5′ splice site is essential for splicing, but not for the other function of the 5′ splice site in RNA protection. This hypothesis was substantiated by the observation that in cotransfection experiments with a suppressor U1 snRNA that does not allow hybrid formation in position +1, but in all other positions of the 5′ splice site, the *env* expression from mutant SV E/X tat$^-$ rev$^-$ GTIII was restored (data not shown).

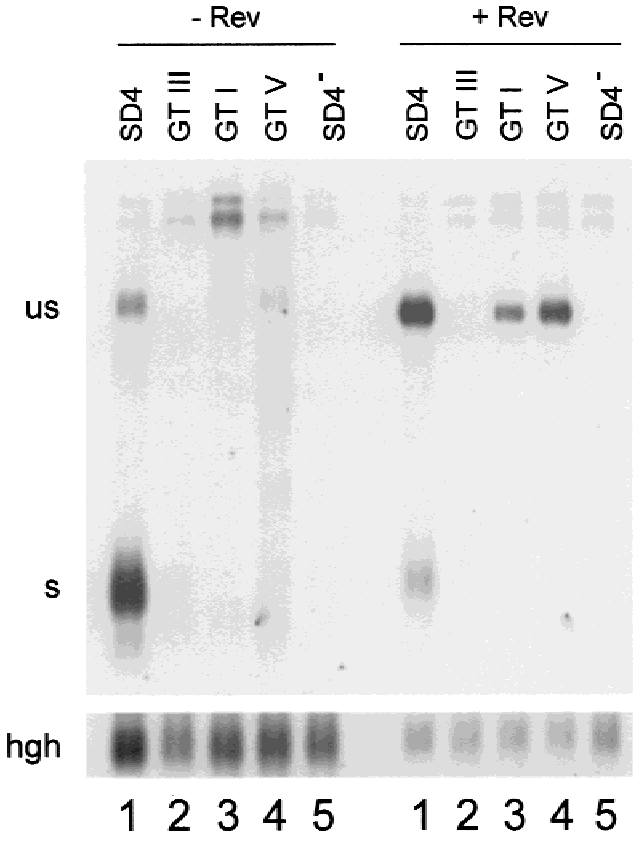

FIGURE 10. Northern blot analysis of poly(A)$^+$ RNA isolated from HeLa-T4$^+$ cells transfected with the *env* expression vector SV E/X tat$^-$ rev$^-$ and the splice function mutants in the presence (+ Rev) and absence (− Rev) of Rev. Lanes 1: SV E/X tat$^-$ rev$^-$; lanes 2: SV E/X tat$^-$ rev$^-$ GTIII; lanes 3: SV E/X tat$^-$ rev$^-$ GTI; lanes 4: SV E/X tat$^-$ rev$^-$ GTV; lanes 5: SV E/X tat$^-$ rev$^-$ SD4$^-$. us: unspliced transcript; s: spliced transcript; hgh: human growth hormone transcript.

Efficient 5′-splice-site-dependent Rev-mediated *env* gene expression requires a sequence upstream of SD4

In the course of studying gene expression, we constructed a variety of *env* expression vectors all carrying

SD4 but with HIV-1 sequences of different length upstream of SD4. Unexpectedly, two of these *env* expression plasmids, which differed only in the presence of 35 nt upstream of SD4 (SVrev, SVsac-sma, Fig. 11A), showed striking differences in *env* expression (Fig. 11B, cf. lanes 3 and 4). The 35-nt region, which is within the coding sequence for *tat* and *rev*, encompasses a purine-rich region, containing several GAR nucleotide motifs, known to be target sequences for SR proteins (Lavigueur et al., 1993; Xu et al., 1993; Elrick et al., 1998). To exclude the possibility that the difference in *env* expression is due to a distance effect, we mutated four of the GAR nucleotide motifs to CAR within the parental construct SV E/X tat$^-$ rev$^-$ (SV E/X tat$^-$ rev$^-$ GAR$^-$), thereby maintaining the distance between the cap structure and SD4 (Fig. 11A). The immunoblot analysis of HeLa-T4$^+$ cells that were transfected with these constructs shows that *env* expression from SV E/X tat$^-$ rev$^-$ GAR$^-$ also is decreased to the same extent as from the deletion construct (SVsac-sma; Fig. 11B, cf. lanes 5 and 4). This decrease in *env* expression was caused by a decrease in the steady-state level of the glycoprotein mRNA as shown by northern blot analysis (Fig. 11C). We can conclude that the mRNA steady-state level was independent of the length of the leader of the glycoprotein mRNA, and that the GAR motifs located 5′ to SD4 are a requirement for an efficient 5′-splice-site-dependent Rev-mediated *env* gene expression.

Mutation of the GAR motifs upstream of SD4 impairs the splicing efficiency and the responsiveness to recombinant SF2/ASF in vitro

The resemblance of the GAR nucleotide motifs with SF2/ASF binding sites in previously characterized exon splicing enhancers (ESE) allowed us to investigate their role in splicing. To this end, we introduced the GAR mutations into SV-SD4/SA7-pA (SV-GAR$^-$ SD4/SA7-pA) and performed an in vitro splicing reaction. Mutating the GAR sequences caused a fourfold decrease in splicing in the presence of exogenously added recombinant SF2/ASF protein as measured by lariat formation, whereas the splicing was highly inefficient in the absence of SF2/ASF independent of the GAR motif (Fig. 12). This suggests that the GAR motifs do function as an ESE, and that the effect is mediated by SF2/ASF, possibly by stabilizing the interaction of U1 snRNP with the SD4 site.

DISCUSSION

Previous transfection experiments with subgenomic HIV-1 expression vectors have led to contradictory conclusions regarding the requirement of a 5′ splice site upstream of the *env* ORF for *env* expression. Several

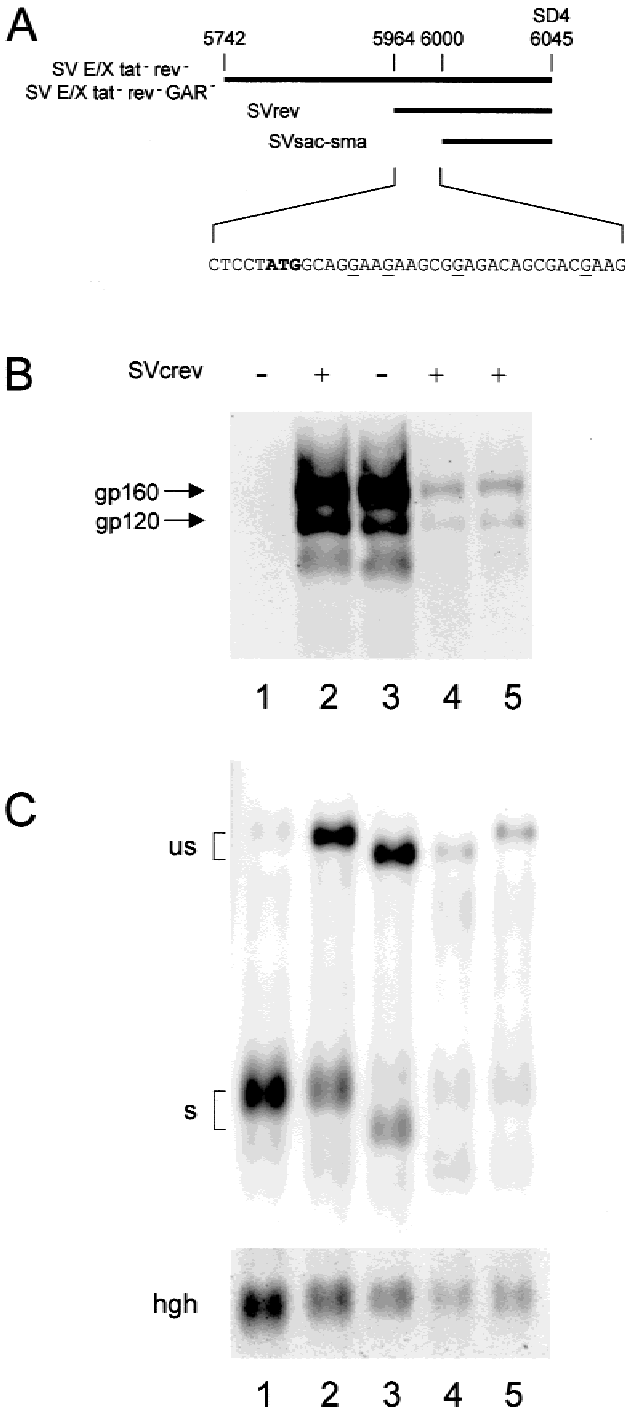

FIGURE 11. **A**: Nucleotide positions of the 5′ ends of the HIV-1 leader sequences present in the *env* expression vectors. The sequence between nt 5964 and nt 6000 is shown beneath. The mutated G residues of the GAR nucleotide motifs within the SV E/X tat$^-$ rev$^-$ GAR$^-$ leader are underlined. The translational start codon of *rev* is indicated in bold. **B**: Immunoblot analysis of glycoproteins expressed in HeLa-T4$^+$ cells transfected with *env* expression vectors carrying different sequences upstream of the *tat*/*rev* 5′ splice site. In addition to 1 μg of *env* expression vector (lanes 1 and 2: SV E/X tat$^-$ rev$^-$; lane 3: SVrev; lane 4: SVsac-sma; lane 5: SV E/X tat$^-$ rev$^-$ GAR$^-$), HeLa-T4$^+$ cells were cotransfected with 1 μg SVcrev (lanes 2, 4, and 5), and 0.5 μg pGL3-control. The positions of gp160 and gp120 are indicated. **C**: Northern blot analysis of poly(A)$^+$ RNA isolated from HeLa-T4$^+$ cells transfected with the *env* expression vectors as indicated in **B**. us: unspliced transcript; s: spliced transcript; hgh: human growth hormone transcript.

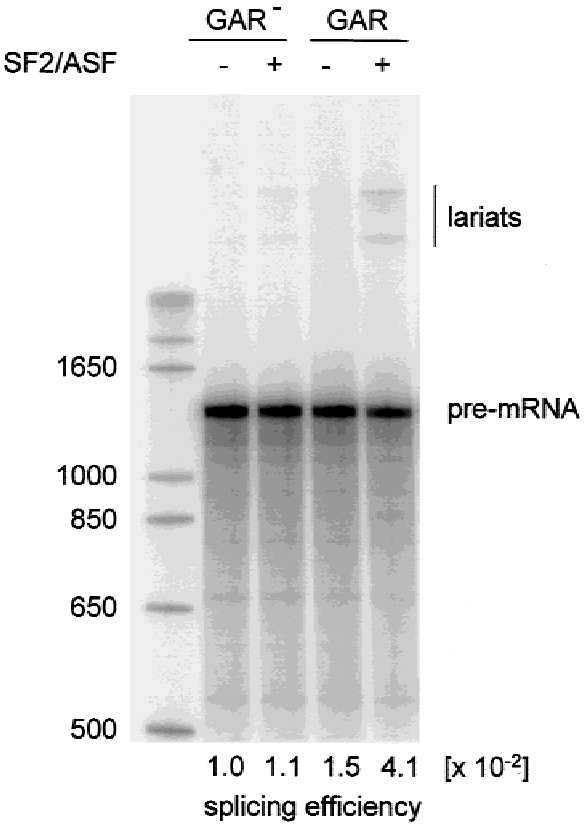

FIGURE 12. In vitro splicing of pre-mRNA produced from SV-SD4/SA7-pA, containing wild-type GAR motifs (GAR), and SV-GAR⁻SD4/SA7-pA, containing four point mutations in the GAR region (GAR⁻) (cf. Fig. 11A). The splicing was performed in the absence and presence of exogenously added recombinant SF2/ASF protein as indicated above the lanes. The splicing efficiency, measured as [lariat RNA species]/[pre-mRNA], was quantitated by PhosphorImager and denoted underneath.

studies have focused on Rev as a prerequisite for splice suppression and nuclear export of *env* mRNAs, with the conclusion that the integrity of the 5′ splice site is necessary for Rev-mediated gene expression (Chang & Sharp, 1989; Lu et al., 1990; Stutz & Rosbash, 1994). In the study by Lu et al. (1990), the participation of the 5′ splice site was established by the demonstration that a mutation in the 5′ splice site in position +5 could be detrimental to *env* expression, but could be offset by a compensatory mutation in an U1 snRNA. It should be pointed out that a mismatch in position +5 should not necessarily be detrimental to *env* expression, because the same +5 mismatch is found in the 5′ splice site SD2 of natural isolates that are able to use SD2 to generate alternatively spliced *env* mRNAs. Others have refuted the need for splice sites in Rev-mediated gene expression based on the successful nuclear export of intronless RRE-containing mRNAs (Fischer et al., 1994; Nasioulas et al., 1994).

In retrospect, these experiments and their interpretation were complicated by different experimental systems and the simultaneous consideration of the effect of splice sites in promoting pre-mRNA stability and Rev response. Our findings indicate that for the intrinsically unstable subgenomic HIV-1 *env* transcripts, the major 5′ splice site SD4 and its hybridization potential to U1 snRNA are essential for prevention of RNA degradation, and that this function is independent of splicing. In addition, we have found that, under our experimental conditions, both the steady-state level of the glycoprotein mRNA and consequently efficient glycoprotein expression requires a purine-rich sequence upstream of the U1 snRNA binding site. This element, which has not been characterized previously, may function by stabilizing U1 snRNP binding. It consists of several GAR motifs similar to consensus binding sequences of ASF/SF2 (Tacke & Manley, 1999) that has been shown to interact with components of U1 snRNP and enhance binding of U1 snRNP to the 5′ splice site (Kohtz et al., 1994; Fig. 11). Our in vitro finding, showing that mutation of the putative SF2/ASF binding site led to a decrease in the efficiency of the splicing reaction in the presence of increased SF2/ASF concentrations, strongly supports this hypothesis.

It may be hypothesized that the observations regarding RNA protection can be extended to other most 5′-located 5′ splice sites. Thus, in the case of HIV-1, it is conceivable that the major 5′ splice site SD1 of the primary mRNA plays a role similar to SD4 in transcripts of subgenomic constructs, based on the observation that, in the genomic RNA, SD1 is in a comparable position relative to the transcriptional start site. If this were the case, any mutation of the major 5′ splice site might not only interfere with the splicing pattern, but also decrease the stabilization of the primary transcript. This would complicate the interpretation of experiments involving mutations of this 5′ splice site, for instance, in the reported splice-site-regulated inactivation of the HIV-1 5′ LTR poly(A) site (Ashe et al., 1995, 1997).

The general observation that many genes require the presence of an intron within the transcription unit of expression vectors to obtain efficient expression has been a matter of discussion since 1979 (Gruss et al., 1979; Hamer & Leder, 1979). Recently, a comparison between the export rates of various pre-mRNAs and their transcribed cDNA counterparts revealed that pre-mRNA splicing can affect the efficiency of mRNA transport (Luo & Reed, 1999). This might include the placement of spliceosome signature proteins at exon–exon junctions as a consequence of splicing. Some proteins, which may be essential components of the nucleocytoplasmic intron-dependent mRNA export pathway, may become associated stably with spliceosome-released mRNA, (Le Hir et al., 2000; Kataoka et al., 2000). In view of our experiments, the stabilizing function of a 5′ splice site may contribute to an increased gene expression prior to and distinct from the splicing step. This is supported by our finding that *env* expression can be detected in the absence of the splice

function of SD4, provided that Rev activates a nucleocytoplasmic export pathway of unspliced mRNA.

The mutational analysis of the 5′ splice site sequence revealed that, in general, positions −3 to +8 can participate in base-pairing to U1 snRNA. The inability of position −4 to compensate for mismatches is compatible with results that show that this position in U1 snRNA is engaged in the internal base pairing within the lower stem of the U1 snRNA (Fig. 3B; Branlant et al., 1981). Lack of conservation at positions +7 and +8 within the pre-mRNA suggests that these positions are not normally participating in the hybridization with the U1 snRNA (Burge et al., 1998). However, these positions can contribute to the formation of the RNA duplex based on the observation that some of our mutants even require an obligatory participation of position +7 and +8 for *env* pre-mRNA protection. A functional implication of these positions has recently been described also for a mutated RSV U1 snRNA binding sequence (Hibbert et al., 1999). Lack of bias towards a consensus nucleotide in these two positions suggests that the central nucleotides of most of the U1 snRNA target sequences are sufficient if not optimal for hybrid formation.

Reliable predictions of U1 snRNA target sequences cannot be based only on sequence compilations of exon/intron borders (Mount, 1982; Padgett et al., 1986; Shapiro & Senapathy, 1987; Burge et al., 1998). Rather, the ability of the sequence to participate in pre-mRNA protection depends more on its capacity to form a hydrogen bonding pattern with the single-stranded 5′ region of U1 snRNA than on the particular nucleotide sequence involved. It remains to be established how important the U1 snRNP-mediated stabilization phenomenon is for expression of cellular genes.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were synthesized and purified as previously described (Schaal et al., 1995). The oligonucleotides are as follows: 135: 5′-GGTGTCGACATAGCAGAATAGGCGTTACTCGACAGAGGAGAGCAAGAA; 136: 5′-GTCTAGGATCTACTGGAGGCCTTTCTTGCTCTCCTCT; 164: 5′-ACAAATAGGGGTTCCGCGCAC; 183: 5′-ATCCTGCAGGGCTTTAGGCTTTGATCCC; 187: 5′-GATGAGCTCTTCGTCGCTGTCTCCGCTTCTTC; 193: 5′-CAAAAGCCTTAGGGATCTCCTACGGCAGGAAGAAGCGGAGACA; 229: 5′-ATCGAATTCTCCTATGGCAGGAAGA; 770 (SD4$^-$): CATCAAGCTTCTCTATCAAAGCACTAATCCGTACCTGTAATGCAACCTATAATAGT; 421 (4C5U): 5′-CATCAAGCTTCTCTATCAAAGCAGTACTTAGTACATGTAATGCA. For the following primers, which are identical to 421 except for the underlined sequence, only these nucleotides are shown; 444 (3U): AGCAGTTAGTAG; 445 (−1G3U): AGCGGTTAGTAG; 446 (5U): AGCAGTAATTAG; 542 (−1G4C8U): AGCGGTACGTAT; 676 (−1G8U): AGCGGTAAGTAT; 753 (−1G5U8U): AGCGGTAATTAT; 759 (−1G5U): AGCGGTAATTAG; 760 (−1G3U8U): AGCGGTTAGTAT; 811 (5U8U): AGCAGTAATTAT; 812 (3U8U): AGCAGTTAGTAT; 813 (4C): AGCAGTACGTAG; 814 (4C8U): AGCAGTACGTAT; 815 (−1G4C): AGCGGTACGTAG; 883 (−1G4C5U8U): AGCGGTACTTAT; 884 (−1G3U7U): AGCGGTTAGTTG; 893 (−2A −1G3U7U): AGAGGTTAGTTG; 900 (4U): AGCAGTATGTAG; 905 (6A7C): AGCAGTAAGACG; 906 (7C): AGCAGTAAGTCG; 1013 (cs–4[16]): CATCCTTAAGCTTCTCTATCAACCAGGTTTCATGTACATGTAATGCA. For the following primers, which are identical to 1013 except for the underlined sequence, only these nucleotides are shown: 1014 (cs − 3[15]): ACAGGTATCATG; 1015 (cs − 2[14]): AGAGGTAACATG; 1018 (cs − 3[13]): ACAGGTTTCATG; 1019 (cs − 4[18]): CCAGGTATCATG; 1001 (GTIV): AGCACTAAGTAG; 1000 (GTIII): AGCACTAAGTAT; 954 (GTI): AGAGCTAAGTAT; 1016 (GTV): ACAGCTAAGTAT; 1171 (αU1 GTIII): CGAAGATCTCATACTTACTGCGCAGGGGAGATACC; 775: 5′-AATTCTGAACACATGTCCGGTTCTAGACCTTAAGATCTTCGAA; 776: 5′-GATCTTCGAAGATCTTAAGGTCTAGAACCGGACATGTGTTCAG; 777: 5′-CTGACATGTCCTGTGGTACCAAGTAGGTTACCGTTACCCGGGTAAACGGATCCAAC; 778: 5′-GCTCTAGACCGTTGGATCCGTTTA; 783: 5′-ATCGGTAACCAGGAGCTGTTGATCC; 784: 5′-AGCGGTACCTGGTGCAGAGAGAAA; 787: 5′-GAGAAGCTTAAGCGTAGTCTGGGACGTCGTATGGGTATTGCTTTGATAGAGATGCTTGATGAGTCTGACT; 800: 5′-TGAGGTTACCGTACTTTCTATAGTGAAT; 825: 5′-ATCCCCGGGGTCTCTGTCTCTCTC; 826: 5′-AGGGAGCTCATCAGAACAGTCAGACTCATCAAG; 937: 5′-AATTCGAGCTCACTCTACCCGGGCAGCT; 938: 5′-GCCCGGGTAGAGTGAGCTCG; 1067 (GAR$^-$): 5′-GATGAGCTCTTGGTCGCTGTCTGCGCTTGTTGCTGCCGTAGGAG; 1172: 5′-GGCAGCAACAAGCGCAGACAGCGACCAAGAGCT; 1173: 5′-CTTGGTCGCTGTCTGCGCTTGTTGCTGC; 1130: 5′-CGAAGATCTCATACTAACCTGGCAGGGGAGAT; 1131: 5′-ATCCTCGAGCCTCCACTGTAGGATTAAC; 1132: 5′-GATCTCCCGGGCTGCAGGATATCCTCGAGTGCA; 1133: 5′-CTCGAGGATATCCTGCAGCCCGGGA.

PCR

PCR amplification of DNA fragments used in cloning procedures was carried out under standard conditions using Pwo DNA polymerase (Roche Molecular Biochemicals). Amplified sequences were verified by automatic sequencing.

Recombinant plasmids

For the parent plasmid SV E/X tat$^-$ rev$^-$, the subgenomic HIV-1 *Eco*RI-*Xho*I fragment of pNLA1 (Strebel et al., 1987), which is a cDNA derivative of pNL4-3 (Adachi et al., 1986), was inserted into a modified multiple cloning site of pSVT7 (Bird et al., 1987) yielding SV E/X. To destroy the translational start codons of *tat* and *rev*, the *Sal*I-*Xho*II fragment (nt 5785–5841; the nucleotides are numbered according to the parent sequence HIV NL43 in the HIV databank of Myers et al. (1995)) was substituted for a PCR-amplified fragment with primers 135 and 136 resulting in an ATG GAG → AGG CCT mutation (SV E/X tat$^-$) and the *Mst*II-*Sac*I (nt 5954–6004) fragment was substituted for a PCR-amplified fragment with primers 193 and 187 resulting in an ATG → ACG

mutation (SV E/X tat$^-$ rev$^-$). For the 5′ splice site mutations, the *Hin*dIII-*Kpn*I fragment (nt 6026–6343) of SV E/X tat$^-$ rev$^-$ was substituted for a PCR product with a 5′ PCR primer carrying the *Hin*dIII-site and the desired mutation of SD4, for example, 444 for the mutation 3U, and the 3′ PCR primer 183. For the construction of SVsac-sma, SVrev, and SV E/X tat$^-$ rev$^-$ GAR$^-$ the *Eco*RI-*Sac*I fragment (nt 5742–6004) of SV E/X tat$^-$ rev$^-$ was substituted for a linker (937/938) and a PCR-amplified fragment with primers 229 and 183 (SVrev) or 164 and 1067 (SV E/X tat$^-$ rev$^-$ GAR$^-$) digested with *Eco*RI and *Sac*I, respectively.

For the construction of the parent splice cassette vector SV-SD4/SA7-pA used in the in vitro splicing reaction, the synthetic poly(A) site of pGL3 (Promega) was inserted upstream of the SV40 early promoter of pSVT7. Subsequently, the *Eco*RI-*Bam*HI fragment of pSVT7 containing the multiple cloning site and the SV40 poly(A) site was substituted for a linker (775/776) containing unique *Afl*III and *Bgl*II sites, which were used to insert at the same time a second linker (777/778) and the SV40 poly(A) site. The following fragments were inserted into the new multiple cloning site: a PCR-amplified *tat* exon 1 *Sac*I/*Hin*dIII fragment with a haemagglutinin-tag at its C terminus (826/787), the 5′ splice site SD4 flanked by unique *Hin*dIII and *Afl*III sites, the *Afl*III-*Kpn*I fragment of pNLA1 (nt 6054–6348) (for SV-SD4/RRE/SA7-pA, the PCR-amplified RRE sequence (nt 7728–7992, 784/783) flanked by *Kpn*I and *Bst*EII), the PCR-amplified 3′ splice site, SA7, (nt 8315–8443, 800/825) flanked by *Bst*EII and *Sma*I, and the *Bam*HI-*Xba*I fragment (nt 8465–8887) of SV E/X tat$^-$ rev$^-$. For the construction of SV-GAR$^-$ SD4/SA7-pA, the *Eco*NI-*Sac*I fragment of SV-SD4/SA7-pA was substituted for an *Eco*NI-*Sac*I adaptor carrying the four GAR $\rightarrow$ CAR point mutations (1172/1173).

For the construction of plasmid pUCBΔU1, the *Bgl*II-*Pst*I fragment (nt 440–726) of the parent plasmid pUCBU1 (kindly provided by M.L. Hammarskjold) was substituted for a linker (1132/1133) containing a unique *Bgl*II, *Pst*I, and *Xho*I site. For the 6A mutation, the *Bgl*II-*Xho*I fragment (nt 440–464) of pUCBΔU1 was substituted for a PCR-amplified fragment (1130/1131), carrying a mutation T $\rightarrow$ A in position 6 of the U1 snRNA. The sequences of all constructs are available on request.

Cell culture and transfection

HeLa-T4$^+$ cells (Maddon et al., 1986) were propagated and transfected for northern blot analysis by electroporation as described previously (Schaal et al., 1995) with minor modifications. Viable cells (3×10^6; trypan blue exclusion) in 200 μL of detach medium (Eagle's minimum essential medium supplemented with 10 mM sodium bicarbonate and 25 mM HEPES/NaOH, pH 7.3) were mixed with CsCl gradient-purified plasmid DNA. The amount of DNA in cotransfection experiments was kept constant by adding pSVT7 (Bird et al., 1987). pXGH5 (Selden et al., 1986) was included in all transfection experiments to monitor transfection efficiency. The transfection volume was adjusted to 350 μL with detach medium. The cells were exposed to a single pulse with a setting of 550 V/cm and 960 μF (Bio-Rad gene pulser with capacitance extender), and the cells were kept at room temperature for 5 min before being transferred to culture flasks containing Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. For western blot analysis, HeLa or HeLa-T4$^+$ cells were transfected with FuGENE^{TM} 6 (Roche Molecular Biochemicals), and the transfection efficiency was monitored by cotransfection of pGL3-control (Promega). The medium was changed after 24 h, and cells were harvested and prepared 48 h after transfection.

Western blot analysis

Cells were scraped from the culture flasks into the medium, sedimented at 12,000 $\times$ *g* for 10 s, washed twice in PBS and suspended in 200 μL of SDS-polyacrylamide gel electrophoresis sample buffer (Laemmli, 1970). An aliquot of the PBS washing step was analyzed for luciferase activity (luciferase assay system, Promega). Protein concentration was measured by a Bradford Protein Assay (Bio-Rad) and adjusted in each sample to equal amounts of luciferase activity and protein amount by adding extracts of mock transfected cells. Samples were subjected to electrophoresis on SDS-7% polyacrylamide gel and transferred to a PVDF-membrane (Immobilon^{TM} P, pore size 0.45 μm; Millipore) by electroblotting with 70 V in transfer buffer (200 mM glycine, 25 mM Tris, 20% methanol) for 1 h. Blots were blocked in PBS with 10% bovine serum albumin (BSA), 10% Tween® 20, for 20 h. Protein detection was performed in PBS, 1% BSA, 1% Tween® 20, with a monoclonal mouse-anti-gp120 antibody (NEA-9305, DuPont, 1:500 or 87-133/026, 1:5,000, kindly provided by Dade Behring) for 1 h, washed four times, incubated with a sheep-anti-mouse antibody conjugated with horseradish peroxidase (NA 931, Amersham, 1:1,000), washed four times, rinsed with water, and visualized by a chemiluminescence detection system (ECL^{TM}-system and ECL^{TM} hyperfilm, Amersham; Super Signal® ultra, Pierce).

Northern blot analysis

Total RNA was isolated 30 h after transfection by the single-step method (Chomczynski & Sacchi, 1987) using RNA-clean (Hybaid-AGS, Heidelberg). The poly(A)$^+$ RNA of 60–85 μg total RNA was isolated with Dynabeads® oligo(dT)$_{25}$ (Dynal, Oslo), subjected to electrophoresis on a 1.2% agarose-1% formaldehyde-gel and blotted onto a positively charged nylon membrane (Roche Molecular Biochemicals). After UV cross-linking (0.5 J/cm^2), the membrane was hybridized with digoxigenin- (DIG) labeled antisense RNA probes in a buffer containing 50% (v/v) formamide, 5$\times$ SSC, 50 mM sodium phosphate, 0.1% (w/v) N-lauroylsarcosine, 7% (w/v) SDS, 2% (w/v) blocking reagent (Roche Molecular Biochemicals), 50 μg yeast RNA/L at 68 °C. To monitor transfection efficiency and RNA loading, the membrane was hybridized with a DIG-labeled antisense RNA probe specific for exon 5 of human growth hormone (hGH) mRNA, which was synthesized from 10 μg of the cotransfected plasmid pXGH5 (Selden et al., 1986). HIV-specific RNA was detected by a DIG-labeled antisense RNA probe specific for the first coding exon of *tat* (nt 5785–5960 *Sal*I to *Sau*I). Hybridization signals were visualized by using an anti-Digoxigenin-AP-Fab fragments (50 mU/mL; Roche Molecular Biochemicals) and chemiluminescence substrate (250 μM CDP-Star^{TM}; Roche Molecular Biochemicals). The blots were exposed to ECL^{TM}-films (Amersham).

In vitro splicing

Radioactive labeled and capped transcripts, used for the in vitro splicing analysis, were prepared as described previously (Kjems et al., 1991). In vitro splicing was performed in a 25 μL reaction volume containing 2.5 μL of 10$\times$ splicing buffer (5 mM ATP, 200 mM creatine phosphate, 25 mM $MgCl_2$, 3.4 mM dithiothreitol, 4 U/μL RNasin (Promega), 2 μL of 400 ng/μL recombinant SF2/ASF protein in SF2/ASF buffer (20 mM HEPES/KOH, pH 7.9, 42 mM $(NH_4)_2SO_4$, 0.5 M GuHCl, 0.5 mM DTT, 0.2 mM EDTA, 15% glycerol), 8 μL of HeLa nuclear extract (Cell Culture Center, Belgium; Dignam et al., 1983) and 2 μL of [^{32}P]-labeled in vitro-transcribed RNA (75,000 cpm), by incubation at 30 °C for 2.5 h. In some reactions, the SF2/ASF proteins were omitted from the SF2/ASF buffer. After phenol/chloroform extraction and ethanol precipitation, samples were resuspended in formamide loading buffer and incubated at 95 °C for 3 min followed by fractionation on a 6% polyacrylamide gel containing 8 M urea, 75 mM Tris-borate, pH 8.3, and 1.5 mM EDTA. Gels were subjected to autoradiography.

ACKNOWLEDGMENTS

We thank K. Strebel and M. Martin for providing plasmid pNLA1, M. Gething for pSVT7, M.L. Hammarskjöld for pUCBU1, R. Axel for HeLa-T4$^+$ cells (through the MRC AIDS Directed Program Reagent Project), and Dade Behring for the anti-gp120 mAb (87-133/026). We also thank Martina Austrup and Maria Thieme for excellent technical assistance and Drs. K. Köhrer and S. Scheuring for plasmid sequencing. This work was supported by the Heinrich-Heine-University, Düsseldorf, (Forschungspool) to H.S., by a grant from the Heinz-Ansmann-Stiftung and a short-term fellowship of the Boehringer Ingelheim Fonds to S.K., T.Ø.T., and M.K.L. were supported by the University of Aarhus.

Received July 10, 2000; returned for revision August 18, 2000; revised manuscript received December 4, 2000

REFERENCES

Adachi A, Gendelman HE, Koenig S, Folks T, Willey R, Rabson A, Martin MA. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol 59*:284–291.

Amendt BA, Si ZH, Stoltzfus CM. 1995. Presence of exon splicing silencers within human immunodeficiency virus type 1 tat exon 2 and tat-rev exon 3: Evidence for inhibition mediated by cellular factors. *Mol Cell Biol 15*:4606–4615.

Ashe MP, Griffin P, James W, Proudfoot NJ. 1995. Poly(A) site selection in the HIV-1 provirus: Inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes & Dev 9*:3008–3025.

Ashe MP, Pearson LH, Proudfoot NJ. 1997. The HIV-1 5′ LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J 16*:5752–5763.

Barrett NL, Li X, Carmichael GG. 1995. The sequence and context of the 5′ splice site govern the nuclear stability of polyoma virus late RNAs. *Nucleic Acids Res 23*:4812–4817.

Bird P, Gething MJ, Sambrook J. 1987. Translocation in yeast and mammalian cells: Not all signal sequences are functionally equivalent. *J Cell Biol 105*:2905–2914.

Bodem J, Lochelt M, Yang P, Flugel RM. 1997. Regulation of gene expression by human foamy virus and potentials of foamy viral vectors. *Stem Cells 15* (Suppl 1):141–147.

Branlant C, Krol A, Ebel JP, Gallinaro H, Lazar E, Jacob M. 1981. The conformation of chicken, rat and human U1A RNAs in solution. *Nucleic Acids Res 9*:841–858.

Burge CB, Tuschl T, Sharp PA. 1998. Splicing of precursors to mRNAs by the spliceosomes. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 525–560.

Caputi M, Mayeda A, Krainer AR, Zahler AM. 1999. hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J 18*:4060–4067.

Chang DD, Sharp PA. 1989. Regulation by HIV Rev depends upon recognition of splice sites. *Cell 59*:789–795.

Chomczynski P, Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem 162*:156–159.

Cullen BR. 1992. Mechanism of action of regulatory proteins encoded by complex retroviruses. *Microbiol Rev 56*:375–394.

Del Gatto-Konczak F, Bourgeois CF, Le Guiner C, Kister L, Gesnel MC, Stevenin J, Breathnach R. 2000. The RNA-binding protein TIA-1 is a novel mammalian splicing regulator acting through intron sequences adjacent to a 5′ splice site. *Mol Cell Biol 20*:6287–6299.

Dignam JD, Lebovitz RM, Roeder RG. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res 11*:1475–1489.

Dyhr-Mikkelsen H, Kjems J. 1995. Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript in vitro. *J Biol Chem 270*:24060–24066.

Elrick LL, Humphrey MB, Cooper TA, Berget SM. 1998. A short sequence within two purine-rich enhancers determines 5′ splice site specificity. *Mol Cell Biol 18*:343–352.

Emerman M, Malim MH. 1998. HIV-1 regulatory/accessory genes: Keys to unraveling viral and host cell biology. *Science 280*:1880–1884.

Emerman M, Vazeux R, Peden K. 1989. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. *Cell 57*:1155–1165.

Felber BK, Hadzopoulou Cladaras M, Cladaras C, Copeland T, Pavlakis GN. 1989. rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc Natl Acad Sci USA 86*:1495–1499.

Fischer U, Meyer S, Teufel M, Heckel C, Lührmann R, Rautmann G. 1994. Evidence that HIV-1 Rev directly promotes the nuclear export of unspliced RNA. *EMBO J 13*:4105–4112.

Fischer U, Pollard VW, Lührmann R, Teufel M, Michael MW, Dreyfuss G, Malim MH. 1999. Rev-mediated nuclear export of RNA is dominant over nuclear retention and is coupled to the Ran-GTPase cycle. *Nucleic Acids Res 27*:4128–4134.

Fu XD, Katz RA, Skalka AM, Maniatis T. 1991. The role of branchpoint and 3′-exon sequences in the control of balanced splicing of avian retrovirus RNA. *Genes & Dev 5*:211–220.

Gruss P, Lai CJ, Dhar R, Khoury G. 1979. Splicing as a requirement for biogenesis of functional 16S mRNA of simian virus 40. *Proc Natl Acad Sci USA 76*:4317–4321.

Hamer DH, Leder P. 1979. Splicing and the formation of stable RNA. *Cell 18*:1299–1302.

Hamm J, Mattaj IW. 1990. Monomethylated cap structures facilitate RNA export from the nucleus. *Cell 63*:109–118.

Hibbert CS, Gontarek RR, Beemon KL. 1999. The role of overlapping U1 and U11 5′ splice site sequences in a negative regulator of splicing. *RNA 5*:333–343.

Hope TJ. 1999. The ins and outs of HIV Rev. *Arch Biochem Biophys 365*:186–191.

Huang Y, Carmichael GG. 1996. A suboptimal 5′ splice site is a cis-acting determinant of nuclear export of polyomavirus late mRNAs. *Mol Cell Biol 16*:6046–6054.

Kataoka N, Yong J, Kim VN, Velazquez F, Perkinson RA, Wang F, Dreyfuss G. 2000. Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm (In Process Citation). *Mol Cell 6*:673–682.

Katz RA, Skalka AM. 1990. Control of retroviral RNA splicing through maintenance of suboptimal processing signals. *Mol Cell Biol 10*:696–704.

Kiss-Laszlo Z, Hohn T. 1996. Pararetro- and retrovirus RNA: Splicing and the control of nuclear export. *Trends Microbiol 4*:480–485.
Kjems J, Frankel AD, Sharp PA. 1991. Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. *Cell 67*:169–178.
Kohtz JD, Jamison SF, Will CL, Zuo P, Lührmann R, Garcia-Blanco MA, Manley JL. 1994. Protein–protein interactions and 5′-splice-site recognition in mammalian mRNA precursors. *Nature 368*: 119–124.
Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*:680–685.
Lavigueur A, La Branche H, Kornblihtt AR, Chabot B. 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes & Dev 7*:2405–2417.
Legrain P, Rosbash M. 1989. Some *cis*- and *trans*-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell 57*:573–583.
Le Hir H, Moore MJ, Maquat LE. 2000. Pre-mRNA splicing alters mRNP composition: Evidence for stable association of proteins at exon–exon junctions. *Genes & Dev 14*:1098–1108.
Lu XB, Heimer J, Rekosh D, Hammarskjold ML. 1990. U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus env mRNA that remains unspliced. *Proc Natl Acad Sci USA 87*:7598–7602.
Luo M, Reed R. 1999. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc Natl Acad Sci USA 96*:14937–14942.
Maddon PJ, Dalgleish AG, McDougal JS, Clapham PR, Weiss RA, Axel R. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell 47*: 333–348.
McNally MT, Beemon K. 1992. Intronic sequences and 3′ splice sites control Rous sarcoma virus RNA splicing. *J Virol 66*:6–11.
Moore MJ, Query CC, Sharp PA. 1993. Splicing of precursors to mRNA by the spliceosome. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 303–357.
Mount SM. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res 10*:459–472.
Myers G, Korber B, Hahn BH, Jeang KT, Mellors JW, McCutchan FE, Henderson LE, Pavlakis GN. 1995. *Human retroviruses and AIDS. A compilation and analysis of nucleic acid and amino acid sequences.* Los Alamos, New Mexico: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.
Nasioulas G, Zolotukhin AS, Tabernero C, Solomin L, Cunningham CP, Pavlakis GN, Felber BK. 1994. Elements distinct from human immunodeficiency virus type 1 splice sites are responsible for the Rev dependence of env mRNA. *J Virol 68*:2986–2993.
O'Reilly MM, McNally MT, Beemon KL. 1995. Two strong 5′ splice sites and competing, suboptimal 3′ splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. *Virology 213*:373–385.
Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA. 1986. Splicing of messenger RNA precursors. *Annu Rev Biochem 55*:1119–1150.
Puig O, Gottschalk A, Fabrizio P, Séraphin B. 1999. Interaction of the U1 snRNP with nonconserved intronic sequences affects 5′ splice site selection. *Genes & Dev 13*:569–580.
Reed R. 2000. Mechanisms of fidelity in pre-mRNA splicing. *Curr Opin Cell Biol 12*:340–345.
Schaal H, Klein M, Gehrmann P, Adams O, Scheid A. 1995. Requirement of N-terminal amino acid residues of gp41 for human immunodeficiency virus type 1-mediated cell fusion. *J Virol 69*: 3308–3314.
Schwartz S, Felber BK, Fenyo EM, Pavlakis GN. 1990. Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *J Virol 64*:5448–5456.
Selden RF, Howie KB, Rowe ME, Goodman HM, Moore DD. 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Mol Cell Biol 6*:3173–3179.
Shapiro MB, Senapathy P. 1987. RNA splice junctions of different classes of eukaryotes: Sequence statistics and functional implications in gene expression. *Nucleic Acids Res 15*:7155–7174.
Si ZH, Amendt BA, Stoltzfus CM. 1997. Splicing efficiency of human immunodeficiency virus type 1 tat RNA is determined by both a suboptimal 3′ splice site and a 10 nucleotide exon splicing silencer element located within tat exon. *Nucleic Acids Res 25*: 861–867.
Si ZH, Rauch D, Stoltzfus CM. 1998. The exon splicing silencer in human immunodeficiency virus type 1 Tat exon 3 is bipartite and acts early in spliceosome assembly. *Mol Cell Biol 18*:5404–5413.
Staffa A, Cochrane A. 1994. The tat/rev intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3′ splice site. *J Virol 68*:3071–3079.
Staffa A, Cochrane A. 1995. Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. *Mol Cell Biol 15*:4597–4605.
Strebel K, Daugherty D, Clouse K, Cohen D, Folks T, Martin MA. 1987. The HIV "A" (sor) gene product is essential for virus infectivity. *Nature 328*:728–730.
Stutz F, Rosbash M. 1994. A functional interaction between Rev and yeast pre-mRNA is related to splicing complex formation. *EMBO J 13*:4096–4104.
Tacke R, Manley JL. 1999. Determinants of SR protein specificity. *Curr Opin Cell Biol 11*:358–362.
Wentz MP, Moore BE, Cloyd MW, Berget SM, Donehower LA. 1997. A naturally arising mutation of a potential silencer of exon splicing in human immunodeficiency virus type 1 induces dominant aberrant splicing and arrests virus production. *J Virol 71*:8542–8551.
Xu R, Teng J, Cooper TA. 1993. The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. *Mol Cell Biol 13*:3660–3674.
Zhang D, Rosbash M. 1999. Identification of eight proteins that cross-link to pre-mRNA in the yeast commitment complex. *Genes & Dev 13*:581–592.
Zhang L, Stoltzfus CM. 1995. A suboptimal src 3′ splice site is necessary for efficient replication of Rous sarcoma virus. *Virology 206*:1099–1107.